AI with Super-computed Data for Monte Carlo Earthquake Hazard Classification

Tsuyoshi Ichimura[1,2], Kohei Fujita[1,2], Takuma Yamaguchi[1], Muneo Hori[1,2], Maddegedara Lalith[1,2] and Naonori Ueda[3]

1. Earthquake Research Institute & Department of Civil Engineering, The University of Tokyo 2. Advanced Institute for Computational Science, RIKEN 3. Center for Advanced Intelligence Project, RIKEN

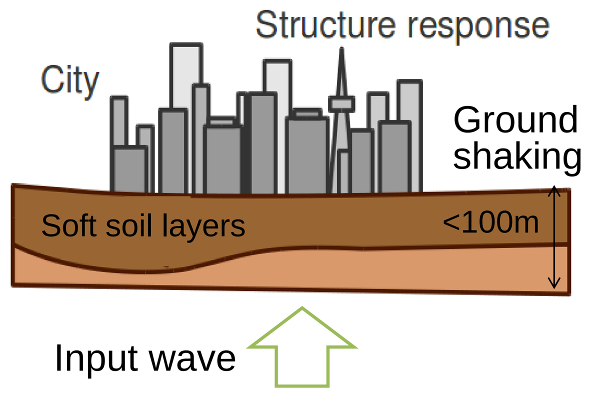
THE UNIVERSITY OF TOKYO RIKEN K computer

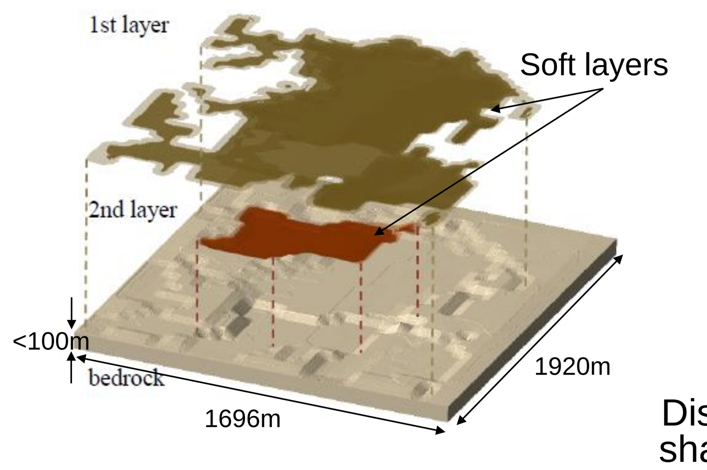
Introduction & Objective

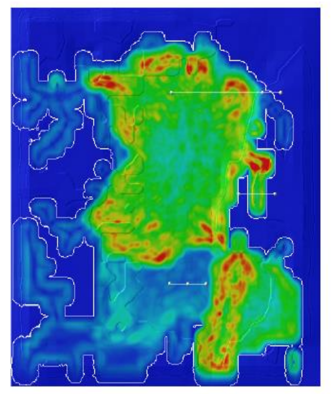
Evaluation of earthquake damage scenarios (i.e., estimation of damage to cities and buildings) required to mitigate from earthquakes

Knowledge of the distribution of shaking (i.e., earthquake hazard) at the ground level in cities with an area of 100km^2 is required to improve damage scenarios

- Soil at a depth of 100 m from the surface has a complex geometry that changes every 10 m
- The ground shaking due to earthquakes becomes highly nonlinear
- Local ground shaking varies significantly with local position [1]







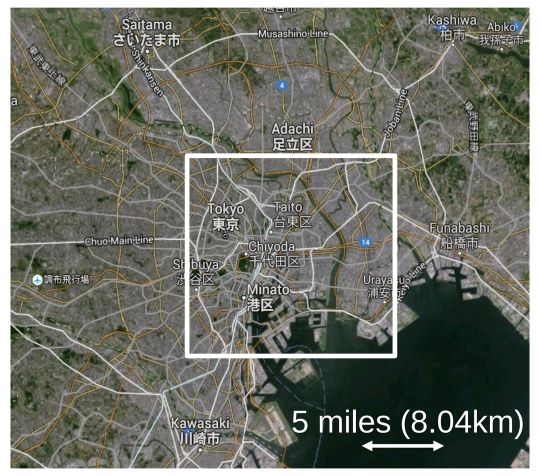
Distribution of earthquake shaking at ground surface

Improvement of hazard maps in higher resolution considering uncertainty

- Observing ground shaking is difficult at high spatial resolution (less than 50 m)
- Simulations using 3D ground information useful for generating hazard map; however, this requires a very large-scale analysis
- Subsurface ground data cannot be measured and its uncertainty must be quantified in the hazard maps
- Monte Carlo simulations suitable for considering uncertainty of data; however, this is impractical because it requires $10^{2\text{-}3}$ repetitions of large-scale analyses

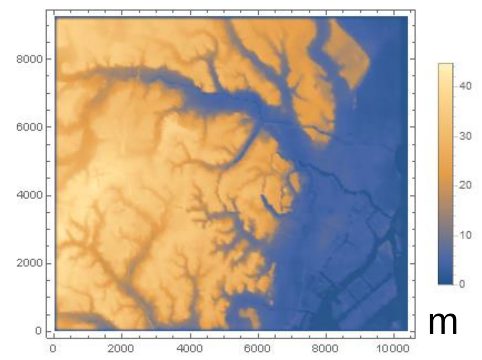
Soil structure data in Tokyo

3D ground information is currently being accumulated for whole urban areas

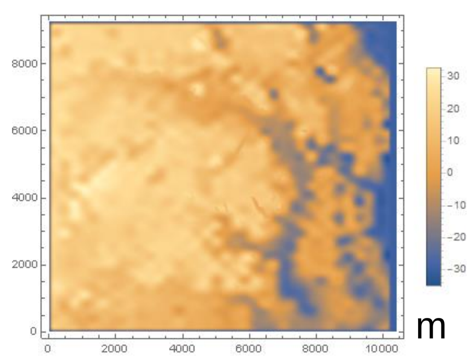


10,250 m × 9,250 m area of Central Tokyo provided by National Digital Soil Map

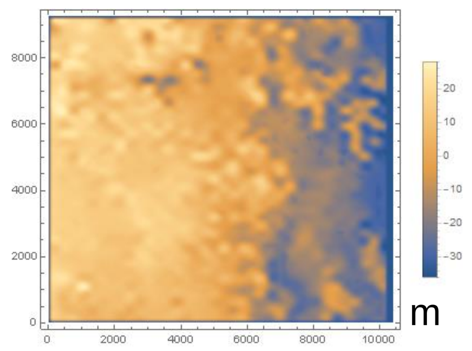
Elevation of subsurface is not highly accurate



Elevation at ground surface



Elevation of interface between 1st and 2nd layer



Elevation of interface between 2nd and 3rd layer

In this study, we develop a practical approach to enable uncertainty analysis of large-scale ground motion analysis

1. *Enable super-large scale analysis by heroic computing using full supercomputer system*
2. *Use super-computed data to construct an AI methodology*
3. *Apply AI methodology to conduct Monte Carlo simulations*

Combination of AI and HPC enables us to step towards high-quality computing (HQC) of Earthquakes

- Verified HPC results itself not sufficient for HQC without considering uncertainty
- Importance of HPC that enables heroic computing is strengthened by combination with AI that makes uncertainty analysis possible

Development of Supercomputing Method

Target problem

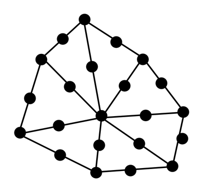
- Implicit nonlinear wave simulation with low-ordered unstructured finite elements
- Much difficult than explicit time integration due to convergence of iterative solver and global communication

Discretized wave equation (**solve for each of 6,000 time-steps**)

$$\left(\frac{4}{dt^2}\mathbf{M}+\frac{2}{dt}\mathbf{C}^n+\mathbf{K}^n\right)\delta\mathbf{u}^n=\mathbf{F}^n-\mathbf{Q}^{n-1}+\mathbf{C}^n\mathbf{v}^{n-1}+\mathbf{M}\left(\mathbf{a}^{n-1}+\frac{4}{dt}\mathbf{v}^{n-1}\right)$$

Matrix (components changes every time step) — Unknown vector (0.1 trillion DOF order) — Known vector

Proposed solver (nominated as SC15 Gordon Bell Prize finalist [2])

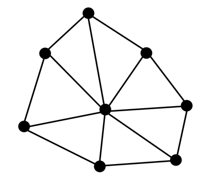


Second ordered tetrahedron

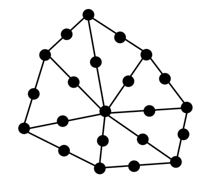
Equation to be solved (double precision)

Outer loop: CG loop — Solving preconditioning matrix; Computations of outer loop

Solving preconditioning matrix:
- Inner coarse loop: Solve system roughly using CG solver (single precision)
- Use as initial solution
- Inner fine loop: Solve system roughly using CG solver (single precision)
- Use for preconditioner of outer loop



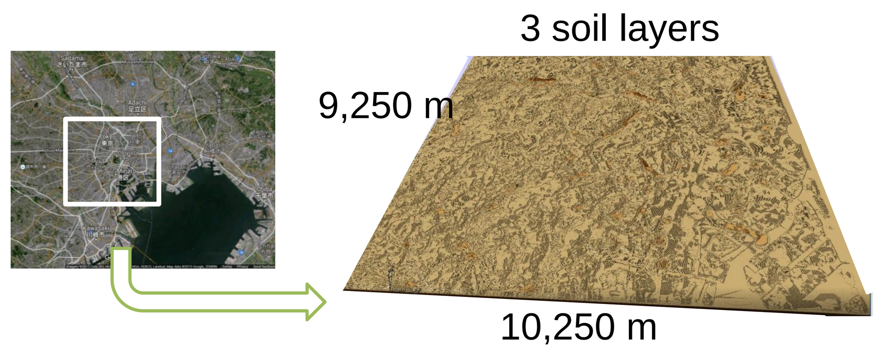
Linear tetrahedron



Second ordered tetrahedron

Structuring unstructured mesh:
Omit costly geometry computation part of unstructured elements

Development of AI with Super-computed Data for Earthquake Hazard Classification

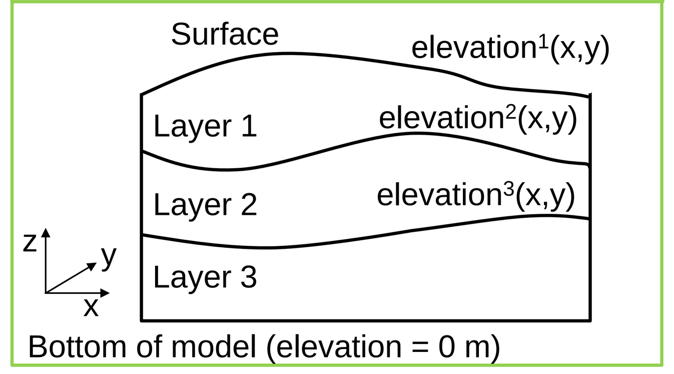


Implicit nonlinear wave simulation in soft soil of Tokyo is conducted by proposed supercomputing method on full K computer (663,552 CPU cores) with max available time (8 hours)

- Minimum element size: 1m (2nd ordered unstructured tetrahedral elements)
- Problem size: 0.133 trillion DOF & 6,600 time-steps
- Analyzed data size: 0.133 trillion DOF × double precision × 6600 time-steps = 6,600 TB

Super-computed ground motion data is classified into 20 levels of severity for 50-m grid points inside an 8,350 m × 7,200 m area of Central Tokyo (24,360 points).

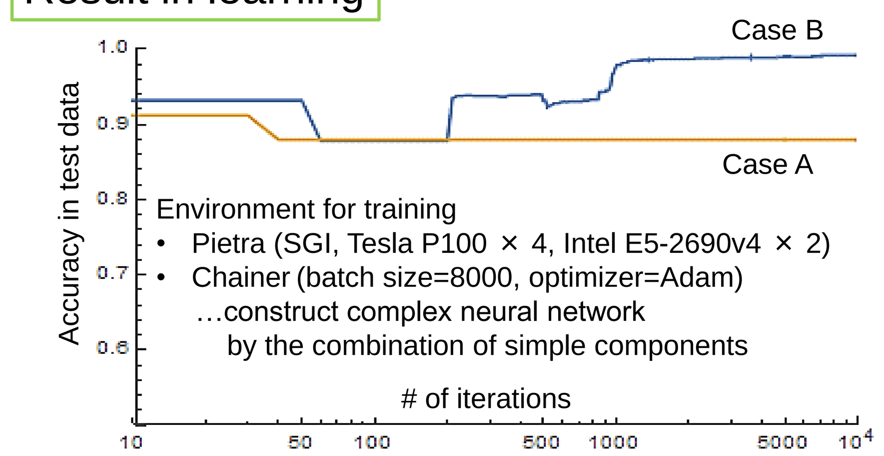
Characteristic value in target problem



Schematic view of AI

Input: elevation data (vector) → Neural Network | Softmax Cross Entropy → Output: estimated earthquake hazard classification (1 scalar)

Result in learning



	Training data	Test data	Check data	Total
# of samples	22360	1000	1000	24360
Case A	19929 (89.1%)	880 (88.0%)	891 (89.1%)	21700 (89.1%)
Case B	22219 (99.4%)	992 (99.2%)	994 (99.4%)	24205 (99.4%)

The number of correct answers and correct ratios. Judged correct if |target value - estimated value| < 2
1000 points are picked up from 23,460 points as test data and check data.

Selection and normalization of input data

Case A: selection and normalization based on seismic design
- Elevation^1(x, y), elevation^2(x, y), elevation^3(x, y) are used as predominant parameters for ground motion at point(x,y) in earthquake resistant designs. These three values are used as input data.

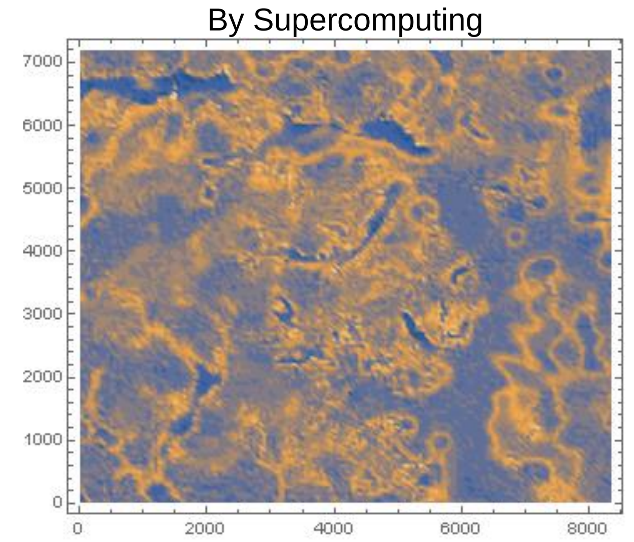
Case B: selection & normalization based on physics
- In the nonlinear wave propagation, the ground motion at point(x, y) is sensitive to the structure within one wavelength of the observation point. Thus 75 points, i.e., elevation^1(x+50 i, y+50 j), elevation^2(x+50 i, y+ 50 j), elevation^3(x+50 i, y+50 j) (i=-2, 2, j=-2, 2) (50m grid) are used as input data.
- To enhance sensitivity, we convert elevation into the thickness of layers and normalized by subtracting average value and divided by standard deviation

Summary: Case B shows better accuracy. We use the AI in Case B in following parts. Case A has lower sensitivity and its accuracy was saturated after 50 iterations. This accuracy was found to be improved to some extent by subtracting a constant and enhancing the sensitivity.

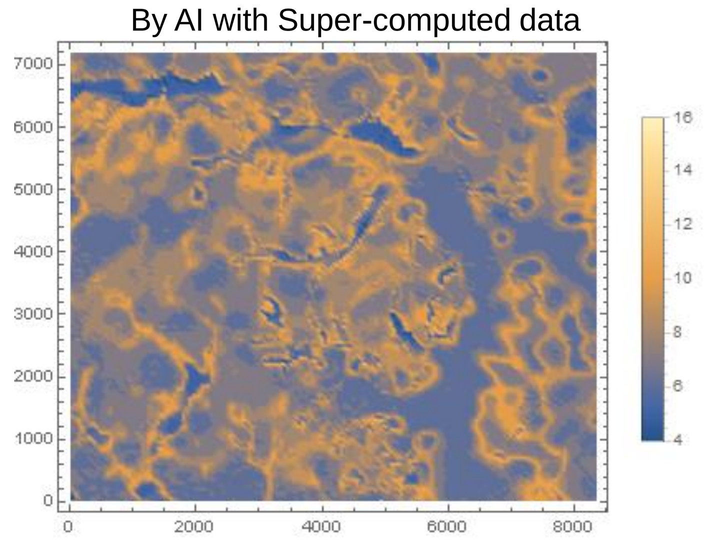
Comparison of earthquake hazard classification

Distribution by AI match adequately to supercomputing results for whole domain

By Supercomputing



By AI with Super-computed data

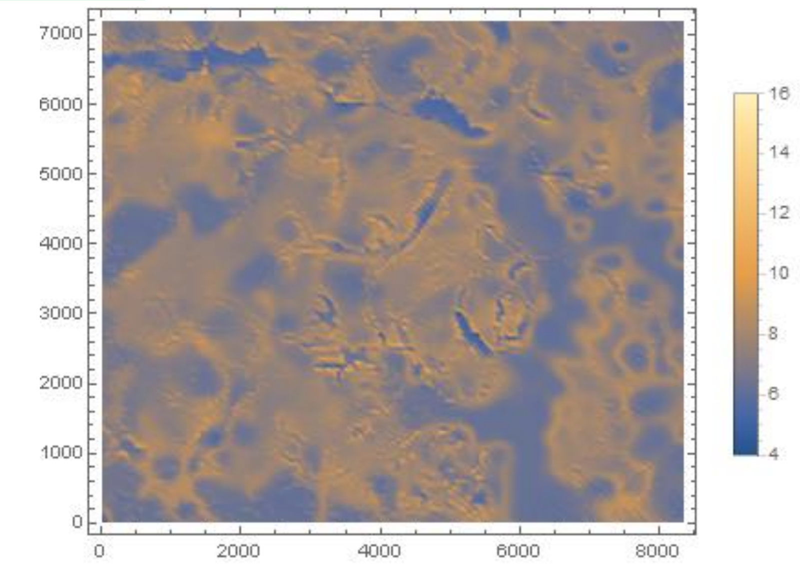


AI-based Monte Carlo Earthquake Hazard Classification

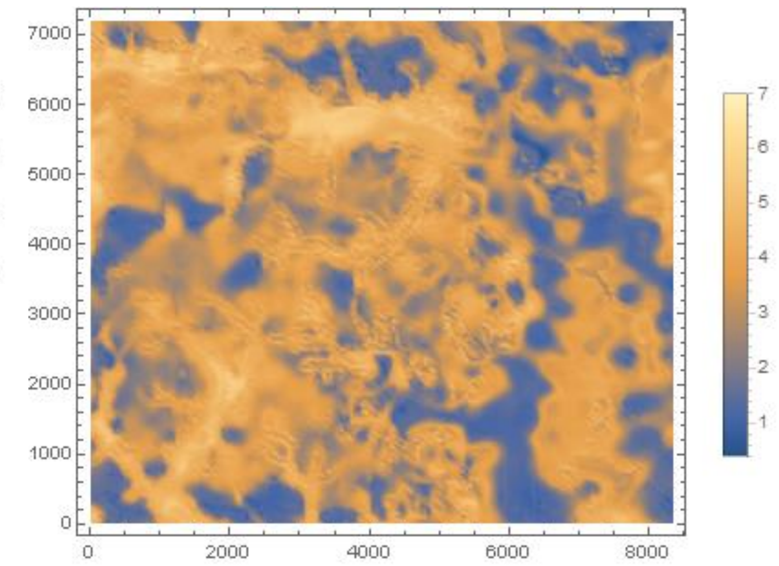
- Change elevation^2(x, y) and elevation^3(x, y) by a normal distribution with a mean of 0 m and a standard deviation of 5 m on a 1,000 m grid
- Conduct a Monte Carlo simulation with 1,000 samples on this stochastic model using the constructed AI

Result

Mean



3 x Standard deviation



- Both mean and standard deviation differ largely from location to location
- This information is first been evaluated via uncertainty analysis using AI

Concluding Remarks

- Highly nonlinear problem with complex resulting distribution and uncertain data targeted: important to consider uncertainty for this problem
- Conducting 1,000 cases of heroic computing is impractical; instead we take a practical approach
- Performed one case of heroic computing using our proposed method and developed an AI methodology based on the obtained super-computed data
- Conducted Monte Carlo Simulation with constructed AI methodology
- High accuracy obtained by selecting and normalizing the training data with consideration of the underlying physical problem
- AI that can confirm permissible level of accuracy can enhance HPC via an uncertainty analysis: large step towards achieving HQC for Earthquakes
- We plan to improve both the AI methodology and the settings for the problem to further improve performance

Acknowledgments

Our results were obtained using the K computer at the RIKEN Advanced Institute for Computational Science (proposal numbers: hp160221, hp160160, hp160157, and hp170249). We acknowledge the support provided by the Japan Society for the Promotion of Science (15K18110, 26249066, 25220908, and 17K14719). We thank the Operations and Computer Technologies Division of RIKEN AICS and the High Performance Computing Infrastructure helpdesk for their support in the use of the K computer.

References

[1] T. Ichimura et al. American Society of Mechanical Engineers, 2014. [2] T. Ichimura et al. SC'15, 2015.